

Weekly Report

Period: 2016/6/20-2016/6/26

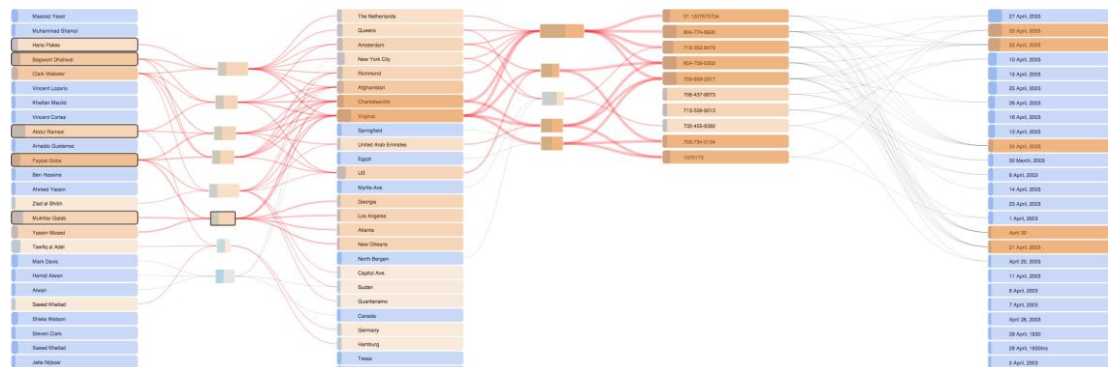
Reporter: Li Zongzhuang

BiSet: Semantic Edge Bundling with Biclusters for Sensemaking

(from TVCG 2016 ,**Author** Maoyuan Sun and so on)

BiSet, a visual analytics technique. It bundles edges based on biclusters to reveal task-oriented semantic insights about potentially coordinated activities for sensemaking. Those edge bundles the first class objects that enable users to directly manipulate relationships.

It has very useful interactions for users. But, in fact, it can only make sense in some specific scenarios and is hardly used when the biclusters are very huge.



Jambalaya: Interactive visualization to enhance ontology authoring and knowledge acquisition in Protégé

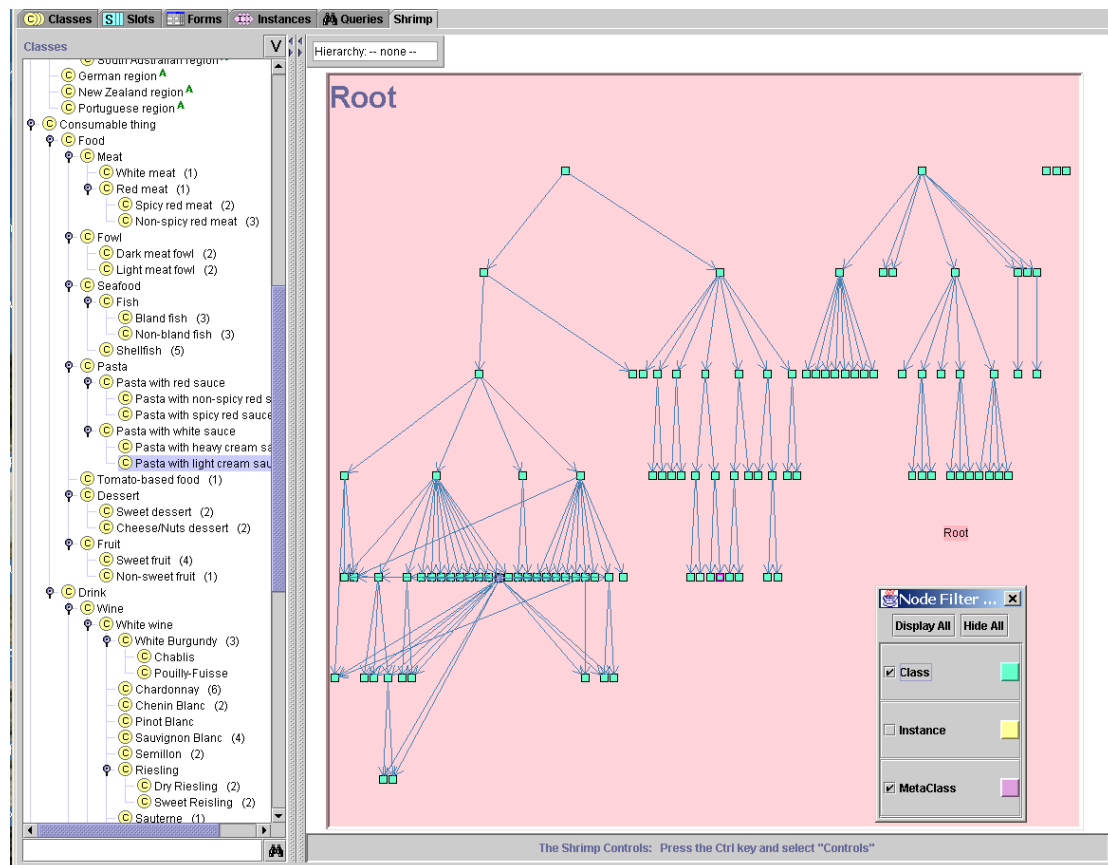
(**Author** Yanhong Wu and so on)

The paper describes how we have applied software visualization techniques to more general knowledge domains.

Jambalaya will result in an easier to use and more powerful environment to support ontology evolution and knowledge acquisition.

In this paper, authors have described the integration of the SHriMP visualization tool which has been used primarily for software visualization with the Protégé tool.

In fact, It a combination about visualization tool and the ontology evolution. We can explore more parts with those techniques.



I See You: Person-of-Interest Search in Social Networks (from ACM ,Author Hsun-Ping Hsieh and so on)

A novel search in online social network, Person-of-Interest (POI) Search, which aims to find a list of desired targets based on a set of user-specified query labels that depict the targets. They develop a greedy heuristic graph search algorithm, which finds the target who not only covers the query labels, but also either possesses better social interactions with peers or has higher social proximity towards the user. It can find a list of targets.

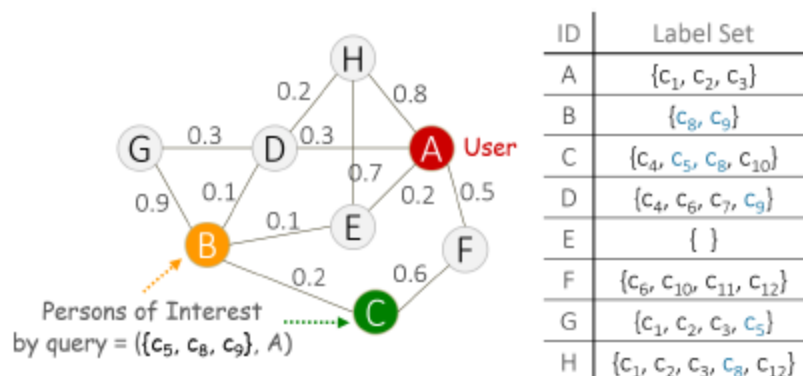


Figure 1: A toy social network to elaborate POI Search, in which each node is associated with a set of labels, and each edge has an interaction cost.

They also make user studies with data from Facebook and Twitter. This is a useful finding which will be used into the realm of heterogeneous information networks.

DBpedia: A Nucleus for a Web of Open Data (from The semantic web 2007

,**Author** Soren Auer and so on)

DBpedia is a community effort to extract structured information from Wikipedia and to make this information available on the Web.

Table 1. The DBpedia datasets

Dataset	Description	Triples
<i>Articles</i>	Descriptions of all 1.95 million concepts within the English Wikipedia including titles, short abstracts, thumbnails and links to the corresponding articles.	7.6M
<i>Ext. Abstracts</i>	Additional, extended English abstracts.	2.1M
<i>Languages</i>	Additional titles, short abstracts and Wikipedia article links in German, French, Spanish, Italian, Portuguese, Polish, Swedish, Dutch, Japanese, Chinese, Russian, Finnish and Norwegian.	5.7M
<i>Lang. Abstracts</i>	Extended abstracts in 13 languages.	1.9M
<i>Infoboxes</i>	Data attributes for concepts that have been extracted from Wikipedia infoboxes.	15.5M
<i>External Links</i>	Links to external web pages about a concept.	1.6M
<i>Article Categories</i>	Links from concepts to categories using SKOS.	5.2M
<i>Categories</i>	Information which concept is a category and how categories are related.	1M
<i>Yago Types</i>	Dataset containing rdf:type Statements for all DBpedia instances using classification from YAGO [16].	1.9 M
<i>Persons</i>	Information about 80,000 persons (date and place of birth etc.) represented using the FOAF vocabulary.	0.5M
<i>Page Links</i>	Internal links between DBpedia instances derived from the internal pagelinks between Wikipedia articles.	62M
<i>RDF Links</i>	Links between DBpedia and Geonames, US Census, Musicbrainz, Project Gutenberg, the DBLP bibliography and the RDF Book Mashup.	180K

This paper describes some emerging applications from the DBpedia community and show how website author scan facilitate DBpedia content within their sites.

In future they will first concentrate on improving the quality of the DBpedia dataset and keep on exploring different types of user interfaces and use cases for the DBpedia datasets.

TripAdvisorN-D: A Tourism-Inspired High-Dimensional Space Exploration

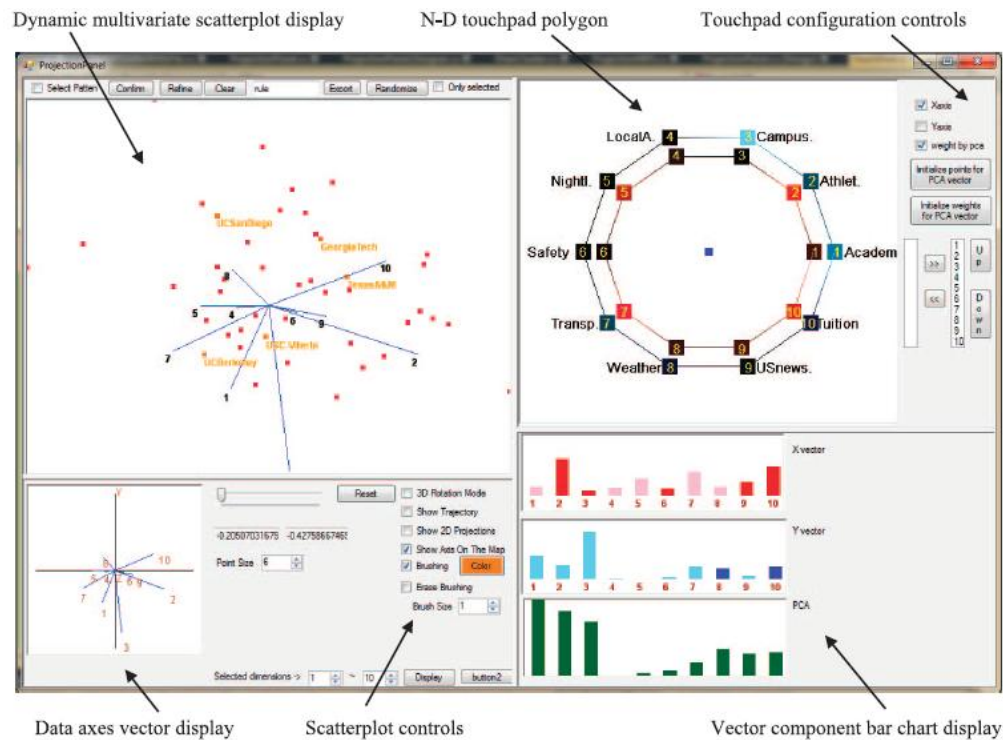
Framework with Overview and Detail (from TVCG2013,**Author** Julia EunJu Nam and so on)

It proposes a multivariate data exploration tool that compares high-dimensional space navigation with a sightseeing trip.

It decomposes this activity into five major tasks:

- 1) Identify the sights
- 2) Plan the trip
- 3) Go on the trip

- 4) Hop off the bus
- 5) Orient and localize



The system within two usage scenarios—a selection task and a clustering task.

The interface consists of two screens: 1) the global sight map 2) the local sight explorer

This interface can be very useful for casual users and Extensive users.

附录：

知识图谱及可视分析

李宗壮

摘要：知识图谱（Knowledge Graph）是谷歌为了提高搜索结果使用语义搜索信息从一个广泛的来源所用的知识库。知识图谱利用了语义的力量，力图通过语义网络的构建来提高搜索的精度和效力。可视化是一种对增强数据认知的技术，它的终极目标是东西蕴含在数据中的现象和规律。知识图谱的可视化，可以有效的提高用户完成搜索目标的效率和精度。数据整合和关联分析推理是可视分析的最佳应用，目前已有不少软件基于这一理念被开发出来。

一、介绍

Google 在 2012 年提出的知识图谱（Knowledge Graph）只要在谷歌搜索引擎里键入单词或短语，在传统搜索结果的右侧就会多出一个栏目直接显示该词条的相关信息，这些信息来自维基百科与其他提供信息服务的网站。它可从许多源获得信息如中情局世界概况、维基数据以及维基百科等。2014 年谷歌公司也推出了新的计划，知识库（Knowledge Vault）。它可以视作知识图谱的潜在继任者。知识库的数据来源比知识图谱更为广阔，它的来源是整个网络，也包括了非结构化的数据源。

视觉是人类获取外部世界信息的最重要通道。可视化，正是对数据进行交互的可视表达

的技术。大数据时代背景下，处理数据的能力已经远远落后于获取数据的能力。而知识图谱所蕴含的数据量无疑是海量的，因此可视化就成为对知识图谱数据处理的一个重要手段。它可以帮助我们更快的恭喜其中的现象和规律并实现目标。然而目前对知识图谱的可视化还比较浅。

二、知识图谱

2.1 相似技术

与知识图谱相关的技术有许多，相似的有例如 DBpedia, Freebase, Knowledge Vault 等等，它们中有些是知识图谱出现的契机，有些是知识图谱的发展。

2.2 语义网

知识图谱的核心理念在于语义的引入，即让计算机可以根据语义判断

三、可视化

可视化是一种方便人进行数据认知的手段。有许多可视化手段被提出。这些都是可以用于知识图谱的表示中的手段

3.1 图数据的可视化

图数据是数据中的一大类。知识图谱的可视化经常以网络或者点线图的形式展现出来

3.2 高维数据的可视化

生活中许多数据的维度都不止一个。因此，高维数据的可视化是常见的可视化类型。知识图谱得到的结果往往是含有多种属性信息的，因此高维数据的可视化也可以给我们以启示。

3.3 其他类型的可视化

因着目标的不同，有许多的可视化方案。其中也会有适合用于知识图谱的可视化方法。

四、应用软件

目前基于人机交互的理论，有许多基于语义网络的软件被提出。该类软件着眼于数据整合和关联分析。其中数据整合主要交由程序后台自主完成，而数据的关联分析则主要依靠人的分析推理能力以及前端的一些交互操作来实现。这就是可视分析的最佳应用。充分发挥人的分析推理特长和计算机的计算特长。

已经有许多致力于这方面的软件如 palantir、IBM i2、Tableau、水晶球等等。